Deep Learning Lecture 1 - Introduction

MA8701 General Statistical Methods

Thiago G. Martins, Department of Mathematical Sciences, NTNU

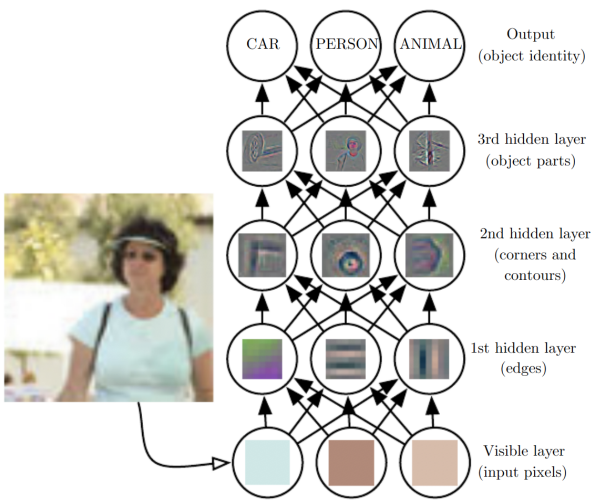
Spring 2019

Main references

- The main references for this lecture is (Chollet and Allaire 2018) and (Goodfellow, Bengio, and Courville 2016).

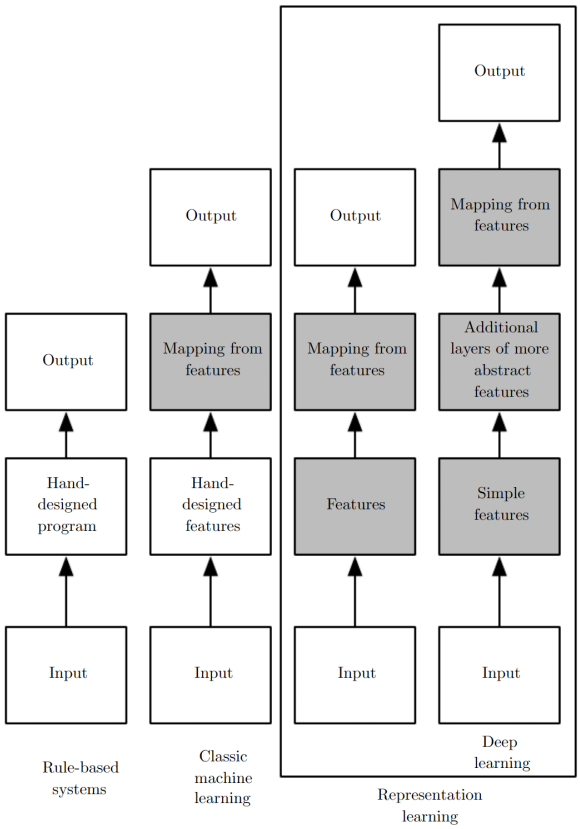
Deep Learning

- Useful tool box to solve intuitive problems.
  - Easy for people, hard to formalise to a machine.
- Useful tool box for automatic feature extraction/representation learning
  - Create representations that are expressed in terms of other, simpler representations.



- Simplified overview of the difference between non-DL and DL models

$$y = f(x; \theta, w) = \phi(x; \theta)^T w$$



Feedforward Networks

- Information flows directly from input to output.
- There are no feedback connections in which outputs of the model are fed back into itself.

Universal approximation property

- Feedforward network with a single layer is sufficient to represent any function.
- However:
  - The layer may be infeasibly large
  - We may fail to learn it.
  - It may fail to generalize correctly.
- In many circumstances, using deeper models can:
  - Reduce the number of units required to represent the desired function.
  - Reduce the amount of generalization error.

Deep Learning and Neuroscience

- Early developments were inspired by neuroscience.
- Neuroscience has now a dimished role in DL research due to our lack of understanding of the brain to a degree that would serve as a guide to us.
- But neuroscience has given us a reason to hope that a single deep learning algorithm can solve many different tasks.

References

Chollet, F., and J. Allaire. 2018. *Deep Learning with R*. Manning Publications. https://books.google.no/books?id=xnIRtAEACAAJ.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press.